

# BUNDESREPUBLIK DEUTSCHLAND



## Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

**Aktenzeichen:** 197 38 205.3

**Anmeldetag:** 2. September 1997

**Anmelder/Inhaber:** Max-Delbrück-Centrum für Molekulare Medizin,  
13125 Berlin/DE

**Bezeichnung:** Mittel zur Diagnose und zur Therapie von Tumorer-  
krankungen

**IPC:** C 07 K, A 61 K, C 12 Q

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ur-  
sprünglichen Unterlagen dieser Patentanmeldung.

**CERTIFIED COPY OF  
PRIORITY DOCUMENT**

München, den 15. Oktober 2004  
Deutsches Patent- und Markenamt  
Der Präsident

Im Auftrag



Max-Delbrück-Centrum für Molekulare Medizin

J. Behrens, W. Birchmeier

## **Mittel zur Diagnose und zur Therapie von Tumorerkrankungen**

### **Zusammenfassung**

Die Erfindung betrifft neue Wege zur Bekämpfung von Tumorerkrankungen durch Ausnutzung molekularbiologischer Zusammenhänge bei der Tumorentstehung.

Der Erfindung liegt die Aufgabe zugrunde, ein Verfahren zur Kontrolle der Regulation von  $\beta$ -Catenin in Körperzellen zu entwickeln.

Gegenstand der Erfindung ist ein neues Protein, welches an  $\beta$ -Catenin bindet und zu dessen zytoplasmatischen Abbau führt. Dieses Protein hat die Aminosäuresequenz gemäß Abb. 1 und wurde als **CONDUCTIN** bezeichnet.

Vom Vorkommen und der Wirkung des Conductins in Körperzellen abgeleitet werden Mittel zur Diagnose und zur Therapie von Tumorerkrankungen entwickelt.

## Beschreibung

Die Erfindung betrifft neue Wege zur Bekämpfung von Tumorerkrankungen durch Ausnutzung molekularbiologischer Zusammenhänge bei der Tumorentstehung. Sie betrifft im einzelnen ein Mittel zur Diagnose von Tumorerkrankungen, und darauf aufbauend ein Mittel zur Therapie. Sie betrifft ferner das neue Protein Conductin, seine Mutanten und Varianten sowie Teile davon, die dazu analogen cDNA-Sequenzen und deren Verwendung in gentherapeutischen und pharmakologischen Verfahren.

Anwendungsgebiete der Erfindung sind die Medizin und die pharmazeutische Industrie.

Cadherine und Catenine bilden Zelladhäsionskomplexe, die in zahlreichen Geweben für die Anheftung der Zellen aneinander verantwortlich sind. Die Cadherine sind Transmembranproteine und stellen den direkten Kontakt zwischen benachbarten Zellen her.  $\alpha$ -,  $\beta$ - und  $\gamma$ -Catenin sind zytoplasmatische Komponenten, die die Cadherine mit dem Aktin-Zytoskelett verbinden. Neben der Funktion bei der Zelladhäsion haben Catenine auch eine entscheidende Rolle bei Signaltransduktionsprozessen.  $\beta$ -Catenin in Vertebraten und das homologe Segmentpolaritäts-Genprodukt Armadillo in Drosophila werden durch den Wnt/Wingless-Signalweg stabilisiert (Nusse, R., Cell 89, 321-323, 1997). Dies führt zu einer Erhöhung der zytoplasmatischen, nicht an Cadherin gebundenen Fraktion dieser Proteine, die daraufhin mit HMG-Transkriptionsfaktoren der LEF-1/TCF-Familie wechselwirken können. Als Resultat wird  $\beta$ -Catenin/Armadillo in den Zellkern transportiert, wo es zusammen mit den LEF/TCF-Proteinen an DNA bindet und bestimmte Gene aktiviert (Behrens, J. et. al., Nature 382, 638-642, 1997).

Dieser Signalweg spielt auch eine Rolle bei der Tumorentstehung. In Kolonepithelzellen wird der zytoplasmatische Pool von  $\beta$ -Catenin durch das Tumorsuppressor-Genprodukt APC (Adenomatosis Polyposis Coli) streng reguliert. Mutationen von APC, wie sie in

etwa 80% aller Kolonkarzinome auftreten, führen zu verkürzten Formen des APC Proteins, die nicht mehr in der Lage sind  $\beta$ -Catenin zu destabilisieren. Dadurch findet man in diesen Tumoren permanente Komplexe von  $\beta$ -Catenin mit dem HMG-Transkriptionsfaktor TCF-4, welche für die Transformation der Zellen verantwortlich gemacht werden. Diese Theorie wird gestützt durch den kürzlichen Befund, daß in Tumoren, in denen APC nicht verändert ist, Mutationen von  $\beta$ -Catenin auftreten. Diese führen ebenfalls zur zytoplasmatischen Stabilisierung von  $\beta$ -Catenin und zur Assoziation mit LEF-1/TCF-Faktoren (Morin, P.J. et. al., Science 275, 1787-1790).

Die Erfindung hat das Ziel, einen neuen Weg zur Verhinderung der Tumorentstehung zu finden. Der Erfindung liegt die Aufgabe zugrunde, ein Verfahren zur Kontrolle der Regulation von  $\beta$ -Catenin in Körperzellen zu entwickeln.

Gegenstand der Erfindung ist ein neues Protein, welches an  $\beta$ -Catenin bindet und zu dessen zytoplasmatischen Abbau führt. Dieses Protein hat die Aminosäuresequenz gemäß Abb. 1 und wurde als CONDUCTIN bezeichnet.

Das Erfindung beruht nun auf der eigenen Erkenntnis, daß Conductin über eine  $\beta$ -Catenin-Bindungsdomäne an  $\beta$ -Catenin und über eine sogenannte RGS-Domäne (Regulator of G-Protein Signalling) an APC-Fragmente bindet. Dadurch kommt es zum zytoplasmatischen Abbau von  $\beta$ -Catenin und in Vertebraten zur Blockade des Wnt/Wingless-Signalwegs. Damit ist klar, daß Conductin ein wichtiger Regulator der  $\beta$ -Catenin-Funktion ist und im Zusammenspiel mit APC zur Tumorsuppression beiträgt.

Davon abgeleitet betrifft die Erfindung ein Mittel zur Diagnose von Tumorerkrankungen, welches dadurch gekennzeichnet ist, daß das Vorhandensein und die Menge von Conductin, seiner Mutanten und Varianten oder seiner Teile in Körperzellen nachgewiesen wird. Dieser Nachweis kann auf der Proteinebene mit spezifischen Antikörpern durchgeführt werden, speziell mit monoklonalen

Antikörpern.

Die Diagnose von Tumorerkrankungen kann gemäß der Erfindung auch auf der Genebene erfolgen. Dazu werden mit ausgewählten Primern und cDNA-Sonden, die aus der Gensequenz des Conductins abgeleitet sind,

- das Gen, das für Conductin, seine Mutanten und Varianten oder Teile davon kodiert, bzw.
- mRNA-Sequenzen, die von diesen Genen abgelesen werden, nachgewiesen.

Das erfindungsgemäße Mittel zur Therapie von Tumorerkrankungen enthält Substanzen, die die Wirkung des Conductins im Körper aktivieren/reaktivieren. Das sind vor allem Mittel, die den Genpromoter des Conductins aktivieren bzw. Mittel, die die Stabilität der von den Conductin-Genen abgeleiteten mRNA-Sequenzen erhöht. Das Hauptziel aller dieser Mittel besteht erfindungsgemäß darin, die Aktivität des Conductins in den Körperzellen zu erhöhen. Dazu kommen u. a. kleinmolekulare Substanzen in Betracht, die z. B. durch High-Throughput-Number-Screening gefunden werden.

Die Erfindung umfaßt auch gentherapeutische Mittel, enthaltend Gene, die für Conductin, seine Mutanten und Varianten oder Teile davon kodieren, bzw. mRNA-Sequenzen, die von diesen Genen abgelesen werden.

Unter Schutz gestellt wird ferner das neue Protein Conductin gemäß Abb. 1, seine Mutanten und Varianten sowie Teile davon. Besonders bevorzugte Teilsequenzen sind die Aminosäuren 78-200 (RGS), 343-464 ( $\beta$ -Catenin-Bindungsdomäne) und 782-832 (Dishevelled Homologie-Region). Zum Schutzmfang gehören auch Teilsequenzen des Adenomatosis Poliposis Coli (APC), gekennzeichnet durch die Aminosäuresequenzen 1464-1604, 1516-1595, 1690-1778 und 1995-2083 als RGS-Domänen-Interaktionsorte.

Gleichermaßen beansprucht werden die analogen cDNA-Sequenzen,

insbesondere die volle cDNA-Sequenz des Conductins (Basenpaare 1-3197) gemäß Abb. 2 sowie die Teilsequenzen des Conductins der Nukleotidfolge 452-820 (RGS-Genabschnitt), der Nukleotidfolge 1247-1612 (Genabschnitt der  $\beta$ -Catenin-Bindungsdomäne) und der Nukleotidfolge 2564-2716 (Genabschnitt der Dishevelled Homologie-Region).

Die Erfindung wird durch die folgenden Ausführungsbeispiele näher erläutert.

Conductin wurde durch einen Hefe 2-Hybrid Screen als  $\beta$ -Catenin-Interaktionspartner identifiziert. Die vollständige cDNA-Sequenz wurde daraufhin isoliert und sequenziert. Die abgeleitete Aminosäuresequenz von Conductin ist in Abb. 1 gezeigt, die Nukleotidsequenz in Abb. 2 und die Gegenüberstellung von Aminosäure- und Nukleotidsequenz in Abb. 3. Conductin besteht aus 839 Aminosäuren und hat ein Molekulargewicht von 92,4 kDa. Durch Sequenzvergleiche wurde im Conductin eine RGS-Domäne (Aminosäuren 78-200) und eine zu dem Protein Dishevelled verwandte Domäne (Aminosäuren 782-832, Dishevelled Homologie-Region) identifiziert (Abb. 1-3). Die  $\beta$ -Catenin-Bindungsdomäne (Aminosäuren 343-464) wurde durch Interaktionsstudien im 2-Hybrid-System entdeckt (Abb. 4). Es zeigte sich, daß diese Domäne ausreichend und notwendig für die Bindung an  $\beta$ -Catenin ist (Abb. 4), wohingegen die RGS- und Dishevelled Homologie-Region nicht beteiligt sind. Die Wechselwirkung von Conductin mit  $\beta$ -Catenin wurde auch in Co-Immunpräzipitationsexperimenten biochemisch bewiesen.

Die Wirkung von Conductin auf  $\beta$ -Catenin wurde in SW480-Zellen untersucht. In diesen Zellen ist das Tumor-Suppressor-Genprodukt APC mutiert, wodurch es zu einem Anstieg des cytoplasmatischen und vor allem nukleären Gehalts von  $\beta$ -Catenin kommt. Die Einbringung von Conductin in diese Zellen führt zu einem drastischen Abbau von  $\beta$ -Catenin, wodurch die Zelle von cytoplasmatischem und im Zellkern befindlichen  $\beta$ -Catenin depletiert wird (Abb. 4). Diese Wirkung auf den Gehalt  $\beta$ -Catenin

ist gleich stark wie die von nichtmutiertem APC, woraus geschlossen werden kann, daß Conductin ebenfalls als Tumorsuppressor durch Regulation von  $\beta$ -Catenin wirkt. Es wurde außerdem gezeigt, daß Conductin den Wnt/Wingless-Signalweg auch in Xenopus-Embryonen durch seine Wirkung auf  $\beta$ -Catenin hemmt.

Es wurde außerdem festgestellt, daß Conductin mit APC direkt interagiert. APC-Fragmente von Aminosäure 1464-1604, 1516-1595, 1690-1778 und 1995-2883 wurden als Interaktionsstellen für Conductin identifiziert. In Conductin erfolgt die Bindung an APC über die RGS-Domäne; dieser Bereich ist ausreichend und notwendig für die Interaktion. Die anderen Domänen in Conductin sind nicht beteiligt (Abb. 5).

Legende zu den Abbildungen:

Abb. 1 Aminosäuresequenz von Conductin

Die Conductin cDNA kodiert ein Protein von 839 Aminosäuren mit einem berechneten Molekulargewicht von 92,4 kDa. Die RGS-Domäne (doppelt unterstrichen), die  $\beta$ -Catenin-Bindungsdomäne (einfach unterstrichen) und die Dishevelled Homologie-Region sind durch Fettdruck hervorgehoben.

Abb. 2 Nukleotidsequenz von Conductin von Position 1-3199

Die Sequenzbereiche sind analog zu Abb.1 markiert.

Abb. 3 Gegenüberstellung von Aminosäure- und Nukleotidsequenz von Conductin

Abb. 4 Analyse der Interaktion von Conductin und seinen Teilen mit  $\beta$ -Catenin

Das Conductin Protein und abgeleitete Teilstücke sind schematisch dargestellt. Hervorgehoben sind die RGS-Domäne (RGS) und die  $\beta$ -Catenin-Bindungsstelle ( $\beta$ -BD). Die Interaktion mit  $\beta$ -Catenin wurde im Hefe 2-Hybrid Assay untersucht und als  $\beta$ -Galaktosidase Einheiten quantifiziert. Man erkennt, daß die Bindung an  $\beta$ -Catenin auf die  $\beta$ -Catenin-Bindungsstelle beschränkt ist, die anderen Teile des Proteins tragen dazu nicht bei. Der Abbau von  $\beta$ -Catenin in SW480 Zellen durch Conductin wurde nach transienter Expression der angegebenen Proteine und Immunfluoreszenz-Färbung von  $\beta$ -Catenin analysiert. Nur Teilstücke von Conductin, die an  $\beta$ -Catenin binden, führen zu dessen Abbau.

Abb. 5 Analyse der Bindung von Conductin und seinen Teilen an APC

Die Bindung der APC Fragmente von Aminosäure 1464-1604 (APCfr.1)

und 1516-1595 (APCfr. 2) an Conductin und abgeleitete Teilstücke wurden im Hefe 2-Hybrid Assay untersucht und als  $\beta$ -Galaktosidase Einheiten quantifiziert. Die Analyse zeigt die ausschließliche Interaktion von APC mit der RGS-Domäne von Conductin. Vergleichbare Ergebnisse für die Bindung an die RGS-Domäne wurden mit APC Fragmenten von Aminosäure 1690-1778 und 1995-2883 erhalten.

### Patentansprüche

1. Mittel zur Diagnose von Tumoren, enthaltend eine Substanz, mit der

- Conductin, seine Mutanten und Varianten oder Teile davon bzw.
- Gene, die für Conductin, seine Mutanten und Varianten oder Teile davon kodieren, bzw.
- m-RNA-Sequenzen, die von diesen Genen abgelesen werden, nachgewiesen werden.

2. Mittel zur Diagnose von Tumoren nach Anspruch 1, enthaltend spezifische Antikörper gegen Conductin, seine Varianten oder Mutanten oder Teile davon.

3. Mittel zur Diagnose von Tumoren nach Anspruch 1 und 2, dadurch gekennzeichnet, daß die spezifischen Antikörper monoklonale Antikörper sind.

4. Mittel zur Diagnose von Tumoren nach Anspruch 1, enthaltend korrespondierende Oligonukleotid-Primer bzw. DNA-Sonden zum Nachweis der Gene und deren Mutationen.

5. Mittel zur Diagnose von Tumoren nach Anspruch 1, enthaltend korrespondierende Oligonukleotid-Primer bzw. DNA-Sonden zum Nachweis der RNA-Sequenzen.

6. Mittel zur Therapie von Tumoren, enthaltend eine Substanz, die die Wirkung des Conductins im Körper aktiviert/reaktiviert.

7. Mittel nach Anspruch 6, enthaltend eine Substanz, die den Genpromoter des Conductins aktiviert.

8. Mittel nach Anspruch 6, enthaltend eine Substanz, die die Stabilität der mRNA-Sequenzen erhöht.

9. Mittel nach Anspruch 6, enthaltend eine Substanz, die die

Aktivität des Conductins erhöht.

10. Conductin, seine Varianten und Mutanten sowie Teile davon.

11. Conductin nach Anspruch 10, gekennzeichnet durch die Aminosäuresequenz 1-839 gemäß Abb. 1, wobei Abb. 1 Bestandteil dieses Anspruchs ist.

12. Teilsequenz des Conductins nach Anspruch 10, gekennzeichnet durch die Aminosäuresequenz 78-200 (RGS-Domäne) der Abb. 1.

13. Teilsequenz des Conductins nach Anspruch 10, gekennzeichnet durch die Aminosäuresequenz 343-464 ( $\beta$ -Catenin-Bindungsdomäne) der Abb. 1.

14. Teilsequenz des Conductins nach Anspruch 10, gekennzeichnet durch die Aminosäuresequenz 782-832 (Dishevelled Homologie-Region) der Abb. 1.

15. Teilsequenzen des Adenomatosis Poliposis Coli (APC), gekennzeichnet durch die Aminosäuresequenzen 1464-1604, 1516-1595, 1690-1778 und 1995-2083 als RGS-Domänen-Interaktionsorte.

16. cDNA-Sequenz von Conductin, seiner Varianten oder Mutanten oder Teilen davon.

17. cDNA-Sequenz des Conductins der Nukleotidfolge 1-3197 der Abb. 2, wobei Abb. 2 Bestandteil dieses Anspruchs ist.

18. cDNA-Teilsequenz des Conductins der Nukleotidfolge 452-820 (RGS-Genabschnitt) der Abb. 2.

19. cDNA-Teilsequenz des Conductins der Nukleotidfolge 1247-1612 (Genabschnitt der  $\beta$ -Catenin-Bindungsdomäne) der Abb. 2.

20. cDNA-Teilsequenz des Conductins der Nukleotidfolge 2564-2716 (Genabschnitt der Dishevelled Homologie-Region) der Abb. 2.

21. Verwendung des Conductin-Gens für die Gentherapie von Tumorerkrankungen, dadurch gekennzeichnet, daß ein Vektor mit dem Conductin-Gen konstruiert wird, anschließend ein Gentransfer in den menschlichen Körper erfolgt und damit die Aktivität des Conductins in Körperzellen wiederhergestellt wird.

MSSAVLVTLPPDSSSFREDAPRPPVPGEEGETPPCQPSVGKVQSTKPMVSSNARRNED 60  
GLGEPEGRASPDSPLTRWTKSLHSLLGDQDGAYLFRTFLEREKCVDTLDFWFACNGFRQM 120  
NLKDTKTLRVAKAIYKRYIENNSVVSQKLPATKTYIRDGIKKQQIGSVMFDQAQTEIOA 180  
VMEENAYQVFLTSIDIYLEYVRSGGENTAYMSNGGLGSLKVLCGYLPTLNEEEWTCADLK 240  
CKLSPTVVGLSSKTLRATASVRSTETAENGFRSFKRSDPVNPYHVGSGYVFAPATSANDS 300  
ELSSDALTDMSMTDSSVDGVPPYRMGSKKQLQREMHRSVKANGQVSLPHFPRTHRLPK 360  
EMTPVEPAFAAELISRLEKIKLELESRHSLEERLQQIREDEEKEGSEQALSSRDGAPVQ 420  
HPLALLPPAAMKRTHKPFWTTSPGSSRPPAVNPLVWVAIAHGPAPPNTTTTTTISSVI 480  
PFRLGASC PVAACPLGGKSFLTKQTTKHHYIHHHAVPKTKEEIEAEATQRVRCLC 540  
PGGTDYCYSKCKSHPKAPEPLPGEQFCGSRGGLPKRNAKGTEPGLALSARDGGMSSAA 600  
GGPQLPGEQGDRSQDVWQWMLSERQSKSKPHSAQSIRKSYPLESARAAPGERVSRHLL 660  
GASGHSRSVARAHPTQDPAMPLTPNLAQLEEACRRLAEVSKPQKQRCCVASQQRDR 720  
NHSAGQAGASP FANPSLAPEDHKEPKKLASVHALQASELVVTYFFCGEEIPYRRMLKAQ 780  
SLTLGHFKEQLSKKGNYRYYFKKASDEFACGRVFEEIWDDETVLPMYEGRILGKVERID 839

AAATAAGCAGCGTTCGCGATGGATTTCGGGGCCACCCGGAGGCCGAGGCCTCGCTCCC 60  
 CAAAGGAGAGCTTGTAAAGAGAGGAGGCTCACATGAGCCCCCTGCTGACTTAAGAG 120  
 AGACCAAGCGATTGTGAGAGGAACGTGAGACGCTCTCCCTCACCATGAGTAGGCCGCTGTAGT 180  
 CAAAACAAATCCAAACTCAGTGCAGACGCTCTCCCTCACCATGAGTAGGCCGCTGTAGT 240  
 GACTCTCCTCCAGATCCCAGCAGCAGCTCCCGAGGGATGCTCCCGGGCCCCCGTTCC 300  
 GGGAGAAGAAGGGAGACCCACCGTGTAGCCTAGTGTGGCAAGGTCCAGTCCACCAA 360  
 ACCTATGCCGTTCCCTCTAAATGCTAGGCCAATGAAGATGGACTGGGGAGGCCGAGGG 420  
 GCGGGCCTCCCCGATCCCCCTTGACCAGGTGGACCAAGTCTTACACTCCTGTGGGG 480  
TGACCAGGATGGTGCATACCTCTCCGGACTTCTGGAGAGGGAGAAATGTGTGGATAC 540  
GCTGGACTTCTGGTTGTGTTAATGGGTCAGGCAGATGAACCTGAAGGATACCAAAAC 600  
TTTGCAGTGGCCAAAGCAATCTATAAGAGGTACATTGAGAACAAACAGCCTGTCTCCAA 660  
GCAGCTGAAGCCCCACCAAGACCTACATACGAGATGGCATCAAGAACAGATCGG 720  
CTCGGTATGTTGACCAGGCACAGACCGAGATCCAGGCAGTGTGAGGAAATGCCTA 780  
CCAGGTGTTGACTCTGACATTGGAATATGTGAGGAGTGGGGGGAAACAC 840  
 AGCTTACATGAGTAACGGGGACTGGGAGCCTAAAGGTCTATGTGGCTACCTCCCCAC 900  
 CTTGAATGAAGAAGAGGGAGTGGACGTGTGGCACCTCAAGTGCAAACCTCACCACCGT 960  
 GGGTGGCTTGCCAGCAAAACTCTCGGGCCACCGCAGTGTGAGATCCACGGAAACAGC 1020  
 TGAAAACGGATTCAAGGTCTTCAAGAGAAGCGACCCAGTCATCCTTATCACGTAGGTTTC 1080  
 CGGCTATGCTTGCACCAGCCACCAGTGCACGACAGCAGCGAGTTATCCAGCGACGCAC 1140  
 GACCGACGATTCCATGTCATGACGGACAGTAGCGTAGATGGAGTCCCTCTTACCGCAT 1200  
 GGGGAGTAAGAAACAACCTCAGAGAGAGATGCATCGAGTGTGAAGGCCAATGCCAAGT 1260  
 GTCTCTACCTCATTTCCGAGAACCCACCGCCTGGCCCAAGGAGATGACGCCCTGTGGAAACC 1320  
 TGCTGCCTCGCCCGGAGCTCATCTCCAGGCTGGAGAAACTGAAAAGTGGAGCTGGAAAG 1380  
 CCCCCATAGTCTGGAGGAGCGGCTGCAGCAGATCCGGGAGGATGAAGAAAAGGAGGGTC 1440  
 TGAGCAGGCCCTGAGCTCACGGGATGGAGCACCGGTCAGCACCCCTGGCCCTCCTACC 1500  
 TCCGGCAGCTATGAAGAGGACCCACAAACCATTTGGACGACCCACCTCTCCAGGTCCTC 1560  
AAGACCCCCGGCTGTCAATCCCCGGCTGGTGTGGTCGCTATAGCCACGGTCCCGCTCCCC 1620  
 GACCAACACCACGACCAACACCACATCAGCAGTGTCAATCCCCGGCTGGAGAAACTGAAAAGTGGAGCTGGAAAG 1680  
 AAGCTGCCCGTGGCTGTTGCCCTCTGGAGGCAAGAGCTTCTGACCAAACAGAC 1740  
 GACGAAGCACGTTCACCAACTACATCCACCAACCGCCGTCCCCAAGACCAAGGAGGA 1800  
 GATCGAGGCAGAACGCCACACAGAGAGTCCGCTGCCTCTGCTCTGGGGAAAGATTATTA 1860  
 TTGCTACTCCAAATGCAAAAGCCACCCGAAGGCTCCAGAGCCCTGCCTGGGGAGCAGTT 1920  
 TTGTGGCAGCAGAGGTGGTACCTGCCAAACGGAATGCAAAGGGCACCGAACCGGGTCT 1980  
 TGCACGTGGCCAGGGATGGAGGGATGTCCAGTGCAGCGGGGGCCCCAGCTTCTGG 2040  
 GGAAGAAGGAGACGGTCACAGGATGTCAGGCTGGAGTGTAGAGAGTGAAGCGGCAGAG 2100  
 CAAGTCCAAGCCCCATAGTGCCAAAGCATAAGAAAGAGCTACCCATTGGAGTCTGCCCG 2160  
 TGCGGCCCCAGGAGAACGAGTCAGCCGGCACCATCTGTTGGGGCCAGCGGACACTCCCG 2220  
 CTCGGTGGCCGGCTACCCATTACCCAGGACCTGCAATGCCCTCCCTACCCAC 2280  
 CAACACTTTGGCACAGTAGAGGAAGCCTGCCGCAGGCTGGCAGAGGTGTCAGGCCA 2340  
 GAAGCAGCGGTGCTGCGTGGCAGTCAGCAGAGGGACAGGAACCAACTCGGCTGTCAG 2400  
 GGCAGGAGCCTCACCCCTGCCAACCAAGCCTGGCTCCAGAAGATCACAAAGAGCCAAA 2460  
 GAAACTGGCAAGTGTCCACCGCCTCCAGGGCAGTGAGCTGGTTGTCACCTACTTTCTG 2520  
 TGGAGAAGAAATTCCATACAGGGAGTGTGAAGGCTAAAGCTTGACCCCTGGGCCACTT 2580  
 CAAGGAGCAGCTCACCAAAAGGGAAATTACAGGTATTATTCAGAAGGCCAGTGACGA 2640  
 ATTGCGCTGGGACGAGTTTGAGGAGATCTGGGACGACGAGACAGTGTCTCCATGTA 2700  
 CGAACGGCAGGATCCGGCAAGTGGAGAGGATCGACTGAGCCTTGGCCTCTCGGCCGT 2760  
 CAACCTGGGCAAGCACCTCGGCCACCATGGAGGCCAAGGCCAGAGACCCCTGTCTCAG 2820  
 GCCTACGGCAAGCACCAAAATTCTGAAGGAAATGAAACCAATTAAAGAACAGAAC 2880  
 CTAGGGAGGGACTGGCCCTGGGCTTCAGGAGGGCGGGGGTATGTTGATCTCAGTCTC 2940  
 CAGGAGCCTGGGTACCGAGATGAGAAAGCCTGAACATTTTATTAAACATGACCACTCTG 3000  
 GGCATAGAAGATGCTAGTGTGTTGAGAGACTGACATACATAATAGATGACTTCTAG 3060  
 GGTCTGAAATTCAAGACTAAGAGAAAAGTGTATAGCTTGGCCGCACAGGAGTCCTT 3120  
 ACTGATATTATTGAACAGTCGATTCCCCCTACCCGCCTCCCTACCCCCCGCCCCGAGT 3180  
 TTATGCTGCTTAAACC 3197

Ab6. 2

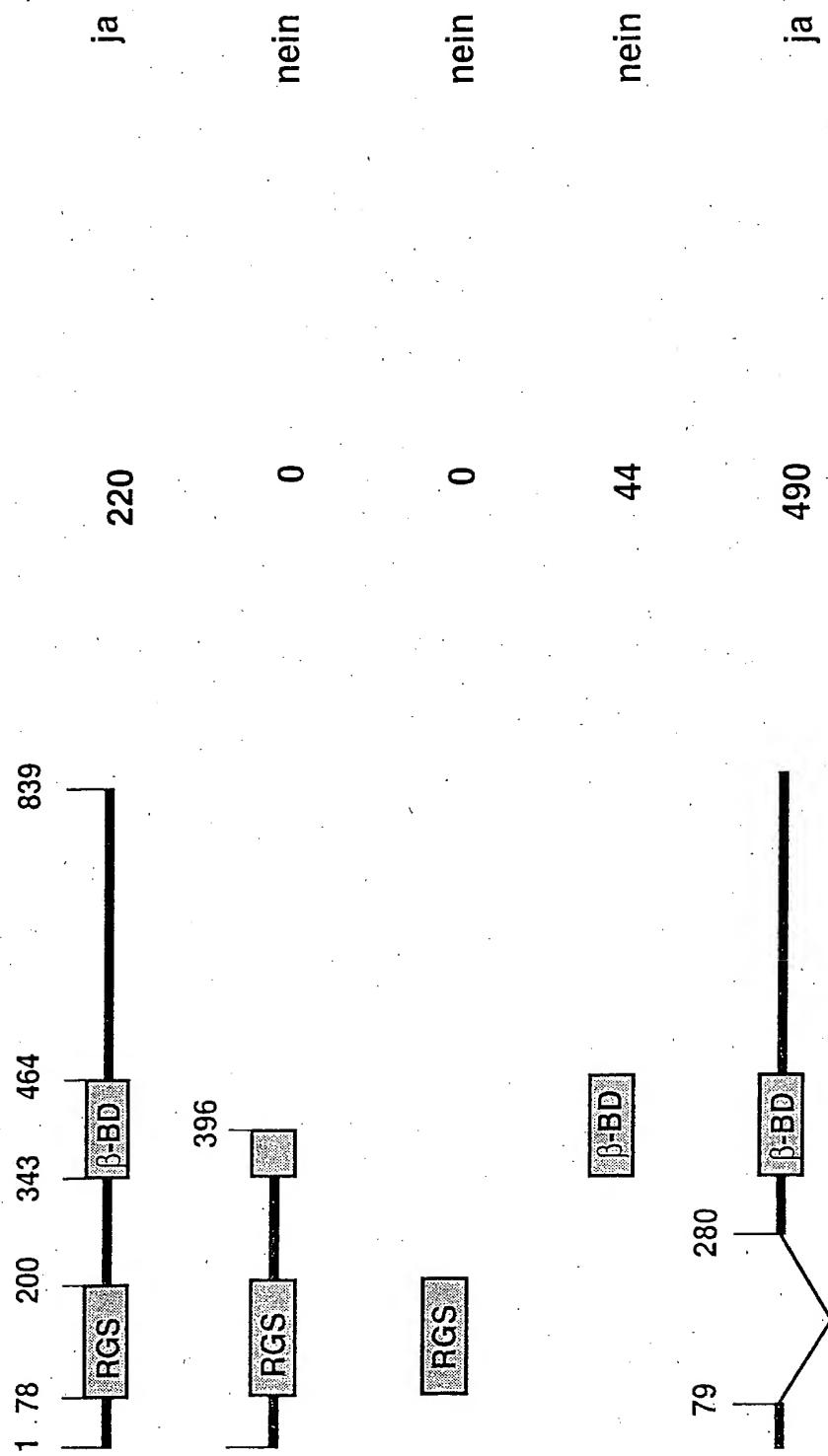
1	5'	-AAATAAGCAGCCGTTCGCGATGGATTTCGGGGCCACCCGGAGGCCGAGGCCTCGCTCCCCAAAGG	66
67		AGAGCTTGTGCTGAAAGAGAGGGAGGCTCACATGAGCCCTGCTGACTTAAGAGAGACCAAGC	129
130		CGATTGCTGAGAGGAAGTGGAAAGAAGAAAAGGAGGGAGGGAAAAAGCAAAACAAATC	192
193		CAAACACTCAGTGAGACGCTCTCCCTCACCATG AGT AGC GCC GTG TTA GTG ACT CTC	247
1		M S S A V L V T L	9
248	CTT	CCA GAT CCC AGC AGC AGC TTC CGC GAG GAT GCT CCG CGG CCC CCG	295
10	L P D P S S S F R E D A P R P P		25
296	GTT	CCG GGA GAA GAA GGG GAG ACC CCA CGG TGT CAG CCT AGT GTG GGC	343
26	V P G E E G E T P P C Q P S V G		41
344	AAG	GTC CAG TCC ACC AAA CCT ATG CCC GTT TCC TCT AAT GCT AGG CGG	391
42	K V Q S T K P M P V S S N A R R		57
392	AAT	GAA GAT GGA CTG GGG GAG CCC GAG GGG CGG GCC TCC CCC GAT TCC	439
58	N E D G L G E P E G R A S P D S		73
440	CCT	TTG ACC AGG TGG ACC AAG TCT TTA CAC TCC TTG TTG GGT GAC CAG	487
74	P L T R W T K S L H S L L G D O		89
488	GAT	GGT GCA TAC CTC TTC CGG ACT TTC CTG GAG AGG GAG AAA TGT GTG	535
90	D G A Y L F R T F L E R E K C V		105
536	GAT	ACG CTG GAC TTC TGG TTT GCT TGT AAT GGG TTC AGG CAG ATG AAC	583
106	D T L D F W F A C N G F R Q M N		121
584	CTG	AAG GAT ACC AAA ACT TTG CGA GTG GCC AAA GCA ATC TAT AAG AGG	631
122	L K D T K T L R V A K A I Y K R		137
632	TAC	ATT GAG AAC AAC AGC GTT GTC TCC AAG CAG CTG AAG CCC GCC ACC	679
138	Y I E N N S V V S K Q L K P A T		153
680	AAG	ACC TAC ATA CGA GAT GGC ATC AAG AAG CAA CAG ATC GGC TCG GTC	727
154	K T Y I R D G I K K O O I G S V		169
728	ATG	TTT GAC CAG GCA CAG ACC GAG ATC CAG GCA GTG ATG GAG GAA AAT	775
170	M F D O A Q T E I O A V M E E N		185
776	GCC	TAC CAG GTG TTC TTG ACT TCT GAC ATT TAC CTG GAA TAT GTG AGG	823
186	A Y Q V F L T S D I Y L E Y V R		201
824	AGT	GGG GGG GAA AAC ACA GCT TAC ATG AGT AAC GGG GGA CTG GGG AGC	871
202	S G G E N T A Y M S N G G L G S		217
872	CTA	AAG GTC TTA TGT GGC TAC CTC CCC ACC TTG AAT GAA GAA GAG GAG	919
218	L K V L C G Y L P T L N E E E E		233
920	TGG	ACG TGT GCC GAC CTC AAG TGC AAA CTC TCA CCC ACC GTG GTT GGC	967
234	W T C A D L K C K L S P T V V G		249
968	TTG	TCC AGC AAA ACT CTT CGG GCC ACC GCG AGT GTG AGA TCC ACG GAA	1015
250	L S S K T L R A T A S V R S T E		265
1016	ACA	GCT GAA AAC GGA TTC AGG TCC TTC AAG AGA AGC GAC CCA GTC AAT	1063
266	T A E N G F R S F K R S D P V N		281
1064	CCT	TAT CAC GTA GGT TCC GGC TAT GTC TTT GCA CCA GCC ACC AGT GCC	1111
282	P Y H V G S G Y V F A P A T S A		297
1112	AAC	GAC AGC GAG TTA TCC AGC GAC GCA CTG ACC GAC GAT TCC ATG TCC	1159
298	N D S E L S S D A L T D D S M S		313
1160	ATG	ACG GAC AGT AGC GTA GAT GGA GTC CCT CCT TAC CGC ATG GGG AGT	1207
314	M T D S S V D G V P P Y R M G S		329
1208	AAG	AAA CAA CTC CAG AGA GAG ATG CAT CGC AGT GTG AAG GCC AAT GGC	1255
330	K K Q L Q R E M H R S V K A N G		345
1256	CAA	GTG TCT CTA CCT CAT TTT CCG AGA ACC CAC CGC CTG CCC AAG GAG	1303
346	Q V S L P H F P R T H R L P K E		361
1304	ATG	ACG CCT GTG GAA CCT GCT GCC TTC GCC GAG CTC ATC TCC AGG	1351
362	M T P V E P A A F A A E L I S R		377
1352	CTG	GAG AAA CTG AAA CTG GAG CTG GAA AGC CGC CAT AGT CTG GAG GAG	1399
378	L E K L K L E L E S R H S L E E		393
<i>PstI V</i>			
1400	CGG	CTG CAG CAG ATC CGG GAG GAT GAA GAA AAG GAG GGG TCT GAG CAG	1447
394	R L Q Q I R E D E E K E G S E Q		409
1448	GCC	CTG AGC TCA CGG GAT GGA GCA CCG GTC CAG CAC CCC CTG GCC CTC	1495
410	A L S S R D G A P V Q H P L A L		425
1496	CTA	CCT CCG GCA GCT ATG AAG AGG ACC CAC AAA CCA TTT TGG ACG ACC	1543
426	L P P A A M K R T H K P F W T T		441



## Conductin Konstrukte

### Interaktion mit $\beta$ -Catenin ( $\beta$ -Galactosidase Einheiten)

### Abbau von $\beta$ -Catenin in SW480 Zellen



16

Abb. 4

10

## Conductin Konstrukte

## Interaktion mit APC-Fragmenten ( $\beta$ -Galactosidase Einheiten)

APC Fr. 1 APC Fr. 2

```

graph TD
    subgraph APC_Fr_1 [APC Fr. 1]
        1[1] --> 78[78]
        78 --> 200[200]
        200 --> 343[343]
        343 --> 464[464]
        464 --> 839[839]
    end
    subgraph APC_Fr_2 [APC Fr. 2]
        839 --> 396[396]
        396 --> 6[6]
        6 --> 9[9]
        9 --> 110[110]
    end
    1 --- RGS1[RGS]
    78 --- RGS1
    200 --- RGS1
    343 --- RGS1
    464 --- RGS1
    839 --- RGS2[RGS]
    396 --- RGS2
    6 --- GBD[G-BD]
    9 --- GBD
    110 --- GBD

```

250

390 0

1

1

6

0  
1  
2

396

RGS

B-BD

17

Ab. 5



Patentanwalt  
Dr. F. Baumbach  
Robert-Rössle-Str. 10

D-13125 Berlin

Dr. F. Baumbach, Robert-Rössle-Str. 10, 13125 Berlin

### Certificate

I, Patent Attorney Dipl.-Chem. Dr. Friedrich Baumbach

declare that I am competent in the German and English languages and I do hereby certify, that the annexed document is the best of knowledge and belief true and correct translation of the

DE 197 38 205.3

Declared at Berlin

this 17 day of February 2005

Patent Attorney Dr. F. Baumbach

---

Tel.: +49-30-94892273  
+49-30-94892274  
Fax: +49-30-94892271

Bankverbindung : Berliner Sparkasse  
BLZ: 10050000  
Kto.-Nr. 1953238820

Max-Delbrück-Centrum für Molekulare Medizin

J. Behrens, W. Birchmeier

**Agent for Diagnosing and Treating Tumor Illnesses**

**Summary**

The invention relates to a new method for combating tumor illnesses through the use of molecular biological associations during formation of the tumor. The aim of the invention is to develop a method for controlling the regulation of beta -catenine in body cells.

The object of the invention is a new protein which bonds to  $\beta$ -catenine and the subsequent cytoplasmic decomposition of said protein. This protein has the amino-acid sequence according to figure 1 and is designated as **conductin**.

Agents for diagnosing and treating tumor illnesses are developed from the occurrence and the action of conductin in body cells.

The invention relates to a new way of combating tumor diseases by utilizing molecular biological relationships of the formation of tumors. In particular, it relates to a material for diagnosing tumor diseases and, based on this, a material for the treatment. It furthermore relates to the new protein, conductin, its mutants and variations as well as parts thereof, to the analogous cDNA sequences and to their use in the gene-therapeutic and pharmacological methods. Areas of the application are medicine and the pharmaceutical industry.

Cadherins and catenins form cell adhesion complexes, which are responsible in numerous tissues for the adhesion of cells to one another. The cadherins are transmembrane proteins and produce the direct contact between adjacent cells.  $\alpha$ -,  $\beta$ - and  $\gamma$ -catenin are cytoplasmic components, which connect the cadherines with the actin cytoskeleton. Aside from their function in cell adhesion, the catenins also play a decisive role in signal transduction processes.  $\beta$ -catenin in vertebrates and the homologous, segment polarity gene product, armadillo in drosophila, are stabilized by the Wnt/wingless signal path (Nusse, R., Cell 89, 321-323, 1997). This leads to an increase in the cytoplasmic fraction of these proteins which is not bound to cadherin, which thereupon could interact with HMG transcription factors of the LEF-1/TCF-family. As a result,  $\beta$ -catenin/armadillo is transported into the cell nucleus where, together with the LEF/TCF proteins, it binds to the DNA and activates certain genes (Behrens, J. et al., Nature 382, 638-642, 1997).

This signal pathway also plays an important role in the formation of tumors. In epithelial cells of the colon, the cytoplasmic pool of  $\beta$ -catenin is strictly regulated by the tumor suppressor gene product APC (Adenomatosis Polyposis Coli). Mutations of APC, such as those occurring in 80% of all colon cancers, lead to shortened forms of APC protein, which are no longer able to destabilize  $\beta$ -catenin. As a result, permanent complexes of  $\beta$ -catenin with the HMG transcription factor TCF-4, which are made responsible for the transformation of the cells, are found in these tumors. This theory is supported by the recent finding that, in tumors in which APC is not changed, mutations of  $\beta$ -catenin occur. These also lead to cytoplasmic stabilization of  $\beta$ -catenin and to an association with the LEF/TCF factors (Morin, P.J. et al., Science 275, 1787-1790).

The invention has the goal of finding a new way for preventing the formation of tumors. It is based on the objective of finding a method for controlling the regulation of  $\beta$ -catenin in cells of the body.

The object of the invention is a new protein, which binds to  $\beta$ -catenin and leads to its cytoplasmic degradation. This protein has the amino acid sequence of Figure 1 and was given the name of conductin.

The invention is based on our own realization that conductin binds to APC fragments over a  $\beta$ -catenin binding domain at  $\beta$ -catenin, and over a so-called RGS-domain (Regulator of G-protein Signaling). As a result, there is cytoplasmic degradation of  $\beta$ -catenin and in vertebrates, blockage of the Wnt/Wingless signal path. With that, it is clear that conductin is an important regulator of the  $\beta$ -catenin function and, in interaction with APC, contributes to the suppression of tumors.

Furthermore, as a consequence, the invention relates to a material for diagnosing tumor diseases, which is characterized in that the presence and the amount of conductin, its mutants and variations or its parts is detected in cells of the body. This detection can be carried out on the protein level with specific antibodies, especially with monoclonal antibodies.

Pursuant to the invention, the diagnosis of tumor diseases can also be carried out on the gene level. For this purpose,

- the gene which codes for conductin, its mutants and variations or parts thereof and/or
- mRNA sequences, which are read off of these genes,

are detected with selected primers and cDNA probes, which are derived from the gene sequence of conductin.

The inventive material for the treatment of tumor diseases contains substances which activate/reactivate the action of the conductin in the body. Above all, these are materials, which activate the gene promoter of conductin or materials, which increase the stability of the mRNA sequences derived from the conductin genes. Pursuant to the invention, the main objective of all these materials consists of increasing the

activity of the conductin in the cells of the body. For this purpose substances of low molecular weight, for example, come into consideration, which are found, for example by high throughput number screening.

The invention also comprises gene therapeutic materials, containing genes, which code for conductin, its mutants and variations of parts thereof, or mRNA sequences, which are read out of these genes.

Furthermore, the new protein conductin of Figure 1, its mutants and variations, as well as parts thereof are placed under protection.

Especially preferred partial sequences are the amino acids 78-200 (RGS), 343-464 ( $\beta$ -catenin-binding domain) and 782-832 (dishevelled homology region). Partial sequences of the Adenomatosis Poliposis Coli (APC), which are characterized by the amino acids 1464-1604, 1516-1595, 1690-1778, 1995-2083 as RGS-domains interaction sites, are also part of the extent of the protection.

Likewise the analogous cDNA sequences, especially the full cDNA sequence of the conductin (base pairs 1-3197) of Figure 2, as well as the partial sequences of the conductin of the nucleotide sequence 452-820 (RGS gene section), of the nucleotide sequence 1247-1612 (gene section of the  $\beta$ -catenin-binding domain) and of the nucleotide sequence 2564-2716 (gene section of the dishevelled homology region).

The invention is explained in greater detail by the following examples.

Conductin was identified by a yeast 2 hybrid screen as a  $\beta$ -catenin interaction partner. The complete cDNA was subsequently isolated and sequenced. The derived amino acid sequence of conductin is shown in Figure 1, the nucleotide sequence in Figure 2 and the gene comparison of the amino acid sequence and the nucleotide sequence is shown in Figure 3. Conductin consists 839 amino acids and has a molecular weight of 92.4 kDa. By a comparison of sequences, an RGS Domain (amino acids 78-200) and a domain (amino acids 782-832, dishevelled homology region), related to the protein dishevelled, were identified (Figures 1-3). The  $\beta$ -catenin-binding domain (amino acids 343-464) was discovered by interaction studies in the 2-hybrid system (Figure 4). It was observed that this domain is

sufficient and necessary for the binding to  $\beta$ -catenin (Figure 4). On the other hand, the RGS homology region and the dishevelled homology region do not participate. The interaction of conductin with  $\beta$ -catenin was also confirmed biochemically in co-immunoprecipitation experiments.

The effect of conductin on  $\beta$ -catenin was investigated in SW480 cells. In these cells, the tumor suppressor gene product APC is mutated, as a result of which there is an increase in the cytoplasmic and especially in the nuclear content of  $\beta$ -catenin. The introduction of conductin into these cells leads to a drastic degradation of  $\beta$ -catenin, as a result of which the cell is depleted of cytoplasmic  $\beta$ -catenin and of  $\beta$ -catenin in the cell nucleus (Figure 4). This effect on the content of  $\beta$ -catenin is equal in intensity to that of not-mutated APC, from which it can be concluded that conductin also acts as a tumor suppressor by regulating  $\beta$ -catenin. Moreover, it was shown that conductin also inhibits the Wnt/wingless signal path in *Xenopus* embryos due to its effect on  $\beta$ -catenin.

Furthermore, it was noted that conductin interacts directly with APC. APC fragments of amino acids 1464-1604, 1516-1595, 1690-1778 and 1995-2883 were identified as interaction sites for conductin. In conductin, the binding to APC takes place over the RGS domains; this region is sufficient and necessary for the interaction. The other domains in conductin do not participate (Figure 5).

### Legends for the figures

#### Figure 1 Amino acid sequence of conductin

The conductin cDNA codes a protein of 839 amino acids with a calculated molecular weight of 92.4 kDa. The RGS domain (double underlining), the  $\beta$ -catenin binding domains (simple underlining) and the dishevelled homology region are emphasized by bold lettering.

#### Figure 2 Nucleotide sequence of conductin at position 1-3199

The sequence regions are marked as in Figure 1.

Figure 3 Comparison of amino amino acid sequence and nucleotide sequence of conductin

Figure 4 Analysis of the interaction of conductin and its parts with  $\beta$ -catenin

The conductin protein and derived partial pieces are shown diagrammatically. The RGS-domain (RGS) and the  $\beta$ -catenin binding site ( $\beta$ -BD) are emphasized. The interaction with  $\beta$ -catenin was investigated in the yeast 2-hybrid assay and quantified as  $\beta$ -galactosidase units. It can be seen that the binding to the  $\beta$ -catenin is limited to the  $\beta$ -catenin binding site; the other parts of the protein do not contribute to this. The degradation of  $\beta$ -catenin in SW480 cells by conductin was analyzed after transient expression of the given proteins and immunofluorescence staining of  $\beta$ -catenin. Only partial pieces of conductin, which bind to  $\beta$ -catenin, lead to this degradation.

Figure 5 Analysis of the binding of conductin and its pieces to APC

The binding of the APC fragments of amino acid 1464-1604 (ApCfr.1) and 1516-1595 (APCfr. 2) to conductin and derived pieces of thereof were analyzed in a 2-hybrid assay and quantified as  $\beta$ -galactosidase units. The analysis reveals the only interaction of APC with the RGS-domain of conductin. Comparable results were derived from APC fragments of amino acid 1690-1778 and 1995-2883.

## Claims

1. A material for the diagnosis of tumors, containing a substance, with which conductin, its mutants and variations or parts thereof or genes, which code for conductin, its mutants and variations or parts thereof or mRNA sequences, which are read off of these genes, are detected.
2. The material for the diagnosis of tumors of claim 1, containing specific antibodies to conductin, its variations or mutants or parts thereof.
3. The material for the diagnosis of tumors of claims 1 and 2, wherein the specific antibodies are monoclonal antibodies.
4. The material for the diagnosis of tumors of claim 1, containing corresponding oligonucleotide primers and/or DNA probes for the detection of the genes and their mutations.
5. The material for the diagnosis of tumors of claim 1, containing corresponding oligonucleotide primers and/or DNA probes for the detection of RNA sequences.
6. A material for the treatment of tumors containing a substance, which activates/reactivates the action of conductin the body.
7. The material of claim 6, containing a substance, which activates the gene promoter of conductin.
8. The material of claim 6, containing a substance, which increases the stability of mRNA sequences.
9. The material of claim 6, containing a substance, which increases the activity of conductin.

10. Conductin, its variations, its mutants as well as parts thereof.
11. The conductin of claim 10, characterized by the amino acid sequence 1-839 of Figure 1, Figure 1 being part of this claim.
12. The partial sequence of conductin of claim 10, characterized by the amino acid sequence 78-200 (RGS-domain) of Figure 1.
13. The partial sequence of conductin of claim 10, characterized by the amino acid sequences 343-464 ( $\beta$ -catenin-binding domain).
14. The partial sequence of conductin of claim 10, characterized by the amino acid sequences 782-832 (dishevelled homology region) of Figure 1.
15. The partial sequences of Adenomatosis Poliposis Coli (APC), characterized by the amino acid sequences 1464-1604, 1516-1595, 1690-1778 and 1995-2083 as the interaction sites of RGS domains.
16. A cDNA sequence of conductin, its variations or mutants or parts thereof.
17. The cDNA sequence of the conductin of the nucleotide sequence 1-3197 of Figure 2, Figure 2 being a component of this claim.
18. The cDNA partial sequence of the conductin of the nucleotide sequence 452-820 (RGS gene section) of Figure 2.
19. The cDNA partial sequence of the conductin of the nucleotide sequence 1247-1612 (gene section of the  $\beta$ -catenin-binding domain) of Figure 2.
20. The cDNA partial sequence of the conductin of the nucleotide sequence 2564-2716 (gene section of the dishevelled homology region) of Figure 2.
21. Use of the conductin gene for the gene therapy of tumor diseases, wherein a vector is constructed with the conductin gene, a gene transfer subsequently

takes place in the human body and, with that, the activity of the conductin in the cells of the body is restored.

MSSAVLVTLLPDPSSSFREDAPRPPVPGEEGETPPCQPSVGKVQSTKPMVSSNARRNED 60  
GLGEPEGRASPDSPLTRWTKSLHSLLGDQDGAYLFRTFLEREKCVDTLDFWACNGFROM 120  
NLKDTKTLRVAKAIYKRYIENNSVVSQKLPATKTYIIRDGIKKQQIGSVMFDQAOETEIQA 180  
VMEENAYQVFLTSIDIYLEYVRSGGENTAYMSNGGLGSLKVLCGYLPTLNEEEWTCADLK 240  
CKLSPTVVGGLSSKTLRATASVRSTETAENGFRSFKRSDPVNPHYVGSGYVFAPATSANDS 300  
ELSSDALTDMSMSMTDSSVDGVPPYRMGSKKQLQREMHSVKANGQVSLPHFPRTHRLPK 360  
EMTPVEPAAFAAELISRLEKLKLESRHSLEERLQOIREDEEKEGSEQALSSRDGAPVQ 420  
HPLALLPPAAMKRTHKPFWTTSPGSSRPPAVNPLVWVIAHGPAPPTTTSTTISSVI 480  
FFFRLGASC PVAACPLLGGKSFLTKQTTKHHVHHYIHHHAVPKTKEEIEAEATQRVRCLC 540  
PGGT DYYCYSKCKSHPKAPEPLPGEQFCGSRGGTLPKRNAKGTEPGLALSARDGGMSSAA 600  
GGPQLPGEEGDRSQDWQWMLSERQSKSPHSAQSIRKSYPLESARAAPGERVSRHLL 660  
GASGHSRSVARAHPTQDPAMPPLTTPNLAQLEEACRRLAEVSKPQKQRCCVASQQRDR 720  
NHSAAGQAGASP FANPSLAPEDHKEPKKLASVHALQASELVVTYFFCGEEIPYRRMLKAQ 780  
SLTLGHFKEQLSKKGNYRYYFKKASDEFACGRVFEEIWDDETVLPMYEGRILGKVERID 839

Figure 1

AAATAAGCAGCCGTCGCGATGGATTTCGGGGCCACCCCGAGGCCGAGGCCTCCGCTCCC 60  
 CAAAGGAGAGCTTGCTGTAAAAGAGAGGAGGCTCACATGAGCCCCCTGCTGACTTAAGAG 120  
 AGACCAAGCCGATTGCTGAGAGGAACCTGAAAGAAGAAAAGGAGGAGGAGGGAAAAAAAG 180  
 CAAAACAAAATCCAAACTCAGTGAGACGCTCTCCCTCACCATGAGTAGCGCCGTGTTAGT 240  
 GACTCTCTTCCAGATCCCAGCAGCAGCTCCCGAGGATGCTCCGCCGCCCCGTTCC 300  
 GGGAGAAGAAGGGAGACCCCACCGTGTAGCCTAGTGAGGGCAAGGTCCAGTCCACCAA 360  
 ACCTATGCCGTTCTCTAATGCTAGGCGGAATGAAGATGGACTGGGGAGCCCGAGGG 420  
 GCGGGCCTCCCCGATTCCCTTGACCAGGTGGACCAAGTCTTACACTCCCTTGGGG 480  
TGACCAGGATGGTGCATACCTCTCCCGACTTCTGGAGAGGGAGAAATGTGTGGATAC 540  
GCTGGACTTCTGGTTGCTGTATGGGTTCAAGGAGATGACCTGAAGGAACCAAAAC 600  
TTTGCAGGTGGCCAAAGCAATCTATAAGAGCTACATTGACAACACAGCTGCTCCAA 660  
GCAGCTGAAGGCCGCCACCAAGACCTACATACCGAGATGGCATCAAGAACAGATCGG 720  
CTCGGTCACTGTTGACCAGGCACAGACCCGACATCCAGGGACTGATGGACCAAAATGCCTA 780  
CCAGGTGTTCTGACTTCTGACATTACCTGAAATATGTGAGGAGTGGGGGAAACAC 840  
AGCTTACATGAGTAACGGGGACTGGGAGCCTAAAGGCTTATGTGCTACCTCCCCAC 900  
CTTGAATGAGAAGAGGAGTGGACGTGTCGGCACCTCAAGTGCACAACTCTCACCCACCGT 960  
GGTTGGCTTGTCGACAAAACCTCTGGGCCACCGCAGGTGAGATCCACGGAAACAGC 1020  
TGAAAACGGATTCAAGGCTCTCAAGAGAACGGACCCAGTCATCCTTATCACGTAGGTT 1080  
CGGCTATGCTTCCACAGGCCACAGTGCACACGACAGCGAGTTATCCAGCGACGCAC 1140  
GACCGACGATTCCATGTCATGACGGACAGTAGCGTAGATGGAGTCCCTTACCGCAT 1200  
GGGGAGTAAGAAACAACCTCCAGAGAGAGATGCATCGCAGTGTGAAGGCCAATGCCAAGT 1260  
GTCTCTACCTCACTTCCGAGAACCCACCGCTGCCAACGACATGACCCCTGTGGAAAC 1320  
TGCTGCCCTCGCCGAGCTCATCTCAGGCTGGAGAAACTGAAACTGGAGCTGCCAAG 1380  
CCGCCATAGTCTGGAGGAGCGGCTGCAGCAGATCCGGAGGATGAAAGAAAAGGAGGGTC 1440  
TGAGCAGGCCCTGACCTCACGGGATGGAGCACCGGTCCAGCACCCCCCTGGCCCTCTTAC 1500  
TCCGGCAGCTAIGAAGAGGACCCACAACCAATTGGAGGACACCCTCTCCAGGGCTC 1560  
AAGACCCCCGGCTGTCATCCCCCTGGTGTGGCTCGTATAGCCCACGGTCCCCTCCCC 1620  
GACCACCAACCAACGACCAACCATCAGCAGTGTCTACCCCTCTTGACTGGGGC 1680  
AAGCTGCCCGTGGCTGCTTGGCCCTCTGGAGGCAAGAGCTTCTGACCAAACAGAC 1740  
GACGAAGCACGTTACCCACCACTACATCCACCAACACGCCGCTCCCAAGACCAAGGAGG 1800  
GATCGAGGCAGAACGCCACACAGAGTCCGCTGCCCTGTCTGGGGAAACAGATTATTA 1860  
TTGCTACTCCAAATGCAAAAGCCACCCGAAGGCTCCAGAGGCCCTGCCCTGGGAGCAGT 1920  
TTGTGGCAGCAGGGTGGTACCTTGCACAAAGGGAATGCAAGAGGCCGAACCGGGTCT 1980  
TGCACTGTGGCCAGGGATGGAGGATGTCCAGTGCAGGGGGGCCCTGGCTTCTGG 2040  
GGAAAGGGAGACCGGTACAGGATGTCTGGCAGTGGATGTTAGAGAGTGAACGGCAGAG 2100  
CAAGTCAAGCCCCATAGTGCACAAAGCATAAGAAAGGCTACCCATTGGAGTCTGCC 2160  
TGCGGCCCCAGGAGAACGAGTCAGCCGGCACCACCTGTTGGGGCCAGCGGACACTCCC 2220  
CTCGGTGCCCGGGCTCACCCATTACCCAGGACCTGCAATGCCCTCCCCCTACCCACC 2280  
CAACACTTGGCACAGCTAGAGGAAGCCTGCCAGGCTGGCAGAGGTGTCGAAGCCCCA 2340  
GAAGCAGGGTGTGCGTGGCCAGTCAGCAGAGGGACAGGAACACTCGGCTGTGGTCA 2400  
GGCAGGAGCCTACCCCTCGCCAACCCAAGCCTGGCTCCAGAAGATCACAAAGGCCAA 2460  
AAAAGTGCAGGTGTCACCGCCTCCAGGCTAGTGACGCTGGTGTCACTTACTTTCTG 2520  
TGGAGAAGAAATTCCATACAGGAGGATGCTGAAGGCTCAAGCTTGAACCCCTGGGCCACTT 2580  
CAAGGAGCAGCTCAGCAAAAGGAAATTACAGGTATTATTCAGAAGAAGGCGAGTGA 2640  
ATTTGCCCTGGGAGGAGTTTGAGGAGAATCTGGGACGACGAGACAGTGTCTCCCATGTA 2700  
CGAAGGCAGGATCTGGCAAAGTGGAGAGGATGACTGAGCCTTGGCCTCTCGGCGTG 2760  
CAACCTGGGCAAGCACCTCGGCGTGCACCATGGAGGCCAGAGGCCAGACCCCTGTCTCAG 2820  
GCCTACGCAACGCCACGAAATATTCTGAAGGAAATGAAACCAATTAGAAAGACAAAGC 2880  
CTAGGGAGGGACTGGCCTGGGCTTCAGGAGGGGGGTATGTTGATCTCAGTCTC 2940  
CAGGAGCCTGGGACCGAGATGAGAACGCCAGTCACTATTATAACATGACCAACTCTG 3000  
GGCTATAGAAGATGCTCAGTGTGTTGAGACACTGACATACATAATAGATGACTTCTAG 3060  
GGTTCTGAATTCTAGACTAAGAGAAAATGCTGTATAGCTTGGCCGACAGGAGTCTT 3120  
ACTGATATTATTGAACAGTCGATTCCCCCTACCCGCTCCCTACCCCCCGCCCCCGAGT 3180  
TTATGCTGCTTAAACC 3197

Figure 2

1	5' -AAATAAGCAGCGGTTCGCGATGGATTGGGGGCCACCCGGAGGGCGTCCGCCCCAAAGG	66
67	AGAGCTTGCTGTAAAGAGAGGGAGGCTCACATGAGCCCTGCTGACTTAAGAGAGACCAAGC	129
130	CGATTGCTGAGAGGAAGTGAAGAAGAAAAGGAGGAGGGAGGGAAAAAGCAAAACAAAATC	192
193	CAAACCTCAGTGAGACGCTCTCCCCTCACCATG AGT AGC GCC GTG TTA GTG ACT CTC	247
1	M S S A V L V T L	9
248	CTT CCA GAT CCC AGC AGC AGC TTC CGC GAG GAT GCT CCG CGG CCC CCG	295
10	L P D P S S F R E D A P R P P	25
296	GTT CCG GGA GAA GAA GGG GAG ACC CCA CCG TGT CAG CCT AGT GTG GGC	343
26	V P G E E G E T P P C Q P S V G	41
344	AAG GTC CAG TCC ACC AAA CCT ATG CCC GTT TCC TCT AAT GCT AGG CGG	391
42	K V Q S T K P M P V S S N A R R	57
392	AAT GAA GAT GGA CTG GGG GAG CCC GAG GGG CGG GCC TCC CCC GAT TCC	439
58	N E D G L G E P E G R A S P D S	73
440	CCT TTG ACC AGG TGG ACC AAG TCT TTA CAC TCC TTG TTG GGT GAC CAG	487
74	P L T R W T K S L H S L L G D Q	89
488	GAT GGT GCA TAC CTC TTC CGG ACT TTC CTG GAG AGG GAG AAA TGT GTG	535
90	D G A Y L F R T F L E R E K C V	105
536	GAT ACG CTG GAC TTC TGG TTT GCT TGT AAT GGG TTC AGG CAG ATG AAC	583
106	D T L D F W F A C N G E R Q M N	121
584	CTG AAG GAT ACC AAA ACT TTG CGA GTG GCC AAA GCA ATC TAT AAG AGG	631
122	L K D T K T L R V A K A I Y K R	137
632	TAC ATT GAG AAC AAC AGC GTT GTC TCC AAG CAG CTG AAG CCC GCC ACC	679
138	Y I E N N S V V S K Q L K P A T	153
680	AAG ACC TAC ATA CGA GAT GGC ATC AAG AAG CAA CAG ATC GGC TCG GTC	727
154	K T Y I R D G I K K Q O I G S V	169
728	ATG TTT GAC CAG GCA CAG ACC GAG ATC CAG GCA GTG ATG GAG GAA AAT	775
170	M F D O A Q T E I Q A V M E E N	185
776	GCC TAC CAG GTG TTC TTG ACT TCT GAC ATT TAC CTG GAA TAT GTG AGG	823
186	A Y Q V F L T S D I Y L E Y V R	201
824	AGT GGG GGG GAA AAC ACA GCT TAC ATG AGT AAC GGG GGA CTG GGG AGC	871
202	S G G E N T A Y M S N G G L G S	217
872	CTA AAG GTC TTA TGT GGC TAC CTC CCC ACC TTG AAT GAA GAA GAG GAG	919
218	L K V L C G Y L P T L N E E E E	233
920	TGG ACG TGT GCC GAC CTC AAG TGC AAA CTC TCA CCC ACC GTG GTT GGC	967
234	W T C A D L K C K L S P T V V G	249
968	TTG TCC AGC AAA ACT CTT CGG GCC ACC GCG AGT GTG AGA TCC ACG GAA	1015
250	L S S K T L R A T A S V R S T E	265
1016	ACA GCT GAA AAC GGA TTC AGG TCC TTC AAG AGA AGC GAC CCA GTC AAT	1063
266	T A E N G F R S F K R S D P V N	281
1064	CCT TAT CAC GTA GGT TCC GGC TAT GTC TTT GCA CCA GCC ACC AGT GCC	1111
282	P Y H V G S G Y V F A P A T S A	297
1112	AAC GAC AGC GAG TTA TCC AGC GAC GCA CTG ACC GAC GAT TCC ATG TCC	1159
298	N D S E L S S D A L T D D S M S	313
1160	ATG ACG GAC AGT AGC GTA GAT GGA GTC CCT CCT TAC CGC ATG GGG AGT	1207
314	M T D S S V D G V P P Y R M G S	329
1208	AAG AAA CAA CTC CAG AGA GAG ATG CAT CGC AGT GTG AAG GCC AAT GGC	1255
330	K K Q L Q R E M H R S V K A N G	345
1256	CAA GTG TCT CTA CCT CAT TTT CCG AGA ACC CAC CGC CTG CCC AAG GAG	1303
346	Q V S L P H F P R T H R L P K E	361
1304	ATG ACG CCT GTG GAA CCT GCT GCC TTC GCC GAG CTC ATC TCC AGG	1351
362	M T P V E P A A F A A E L I S R	377
1352	CTG GAG AAA CTG AAA CTG GAG CTG GAA AGC CGC CAT AGT CTG GAG GAG	1399
378	L E K L K L E L E S R H S L E E	393
PstI V		
1400	CGG CTG CAG CAG ATC CGG GAG GAT GAA GAA AAG GAG GGG TGT GAG CAG	1447
394	R L Q Q I R E D E E K E G S E Q	409
1448	GCC CTG AGC TCA CGG GAT GGA GCA CGG GTC CAG CAC CCC CTG GCC CTC	1495
410	A L S S R D G A P V Q H P L A L	425
1496	CTA CCT CCG GCA CCT ATG AAG AGG ACC CAC AAA CCA TTT TGG ACG ACC	1543
426	L P P A A M K R T H K P F F W T T	441

Figure 3

1544	ACC TCT CCA GGG TCC TCA AGA CCC CCG GCT GTC AAT CCC CTG GTG TGG	1591
442	T S P G S S R P P A V N P L V W	457
1592	GTC GCT ATA GCC CAC GGT CCC GCT CCC CCG ACC ACC ACC ACC AGC ACC	1639
458	V A I A H G R A P P T T T T S T	473
1640	ACC ACC ATC AGC AGT GTC ATA CCC TTC TTT CGA CTG GGG GCA AGC TGC	1687
474	T T I S S V I P F F R L G A S C	489
1688	CCC GTG GCT GTC CCC CTC CTT GGA GGC AAG AGC TTC CTG ACC AAA	1735
490	P V A A C P L L G G K S F L T K	505
1736	CAG ACG ACG AAG CAC GTT CAC CAC TAC ATC CAC CAC CAC GCC GTC	1783
506	Q T T K H V H H Y I H H H A V	521
1784	CCC AAG ACC AAG GAG GAG ATC GAG GCA GAA GCC ACA CAG AGA GTC CGC	1831
522	P K T K E E I E A E A T Q R V R	537
1832	TGC CTC TGT CCT GGG GGA ACA GAT TAT TAT TGC TAC TCC AAA TGC AAA	1879
538	C L C P G G T D Y Y C Y S K C K	553
1880	AGC CAC CCG AAG GCT CCA GAG CCC CTG CCT GGG GAG CAG TTT TGT GGC	1927
554	S H P K A P E P L P G E Q F C G	569
1928	AGC AGA GGT GGT ACC TTG CCA AAA CGG AAT GCA AAG GGC ACC GAA CGC	1975
570	S R G G T L P K R N A K G T E P	585
1976	GGT CTT GCA CTG TCG GCC AGG GAT GGA GGG ATG TCC AGT GCA GCG GGG	2023
586	G L A L S A R D G G M S S A A G	601
2024	GGC CCC CAG CTT CCT GGG GAA GAA GGA GAC CGG TCA CAG GAT GTC TGG	2071
602	G P Q L P G E E G D R S Q D V W	617
2072	CAG TGG ATG TTA GAG AGT GAG CGG CAG AGC AAG TCC AAG CCC CAT AGT	2119
618	Q W M L E S E R Q S K S K P H S	633
2120	GCC CAA AGC ATA AGA AAG AGC TAC CCA TTG GAG TCT GCC CGT GCG GCC	2167
634	A Q S I R K S Y P L E S A R A A	649
2168	CCA GGA GAA CGA GTC AGC CGG CAC CAT CTG TTG GGG GCC AGC GGA CAC	2215
650	P G E R V S R H H L L G A S G H	665
2216	TCC CGC TCG GTG GCC CGG GCT CAC CCA TTT ACC CAG GAC CCT GCA ATG	2263
666	S R S V A R A H P F T Q D P A M	681
2264	CCT CCC CTT ACC CCA CCC AAC ACT TTG GCA CAG CTA GAG GAA GCC TGC	2311
682	P P L T P P N T L A Q L E E A C	697
2312	CGC AGG CTG GCA GAG GTG TCG AAG CCC CAG AAG CAG CGG TGC TGC GTG	2359
698	R R L A E V S K P Q K Q R C C V	713
2360	GCC AGT CAG CAG AGG GAC AGG AAC CAC TCG GCT GCT GGT CAG GCA GGA	2407
714	A S Q Q R D R N H S A A G Q A G	729
2408	GCC TCA CCC TTC GCC AAC CCA AGC CTG GCT CCA GAA GAT CAC AAA GAG	2455
730	A S P F A N P S L A P E D R K E	745
2456	CCA AAG AAA CTG GCA AGT GTC CAC GCG CTC CAG GCC AGT GAG CTG GTT	2503
746	P K K L A S V H A L Q A S E L V	761
2504	GTC ACC TAC TTT TTC TGT GGA GAA GAA ATT CCA TAC AGG AGG ATG CTG	2551
762	V T Y F F C G E E I P Y R R M L	777
2552	AAG GCT CAA AGC TTG ACC CTG GGC CAC TTC AAG GAG CAG CTC AGC AAA	2599
778	K A Q S L T L G H F K E Q L S K	793
2600	AAG GGA AAT TAC AGG TAT TAT TTC AAG AAG GCG AGT GAC GAA TTT GCC	2647
794	K G N Y R Y F K K A S D E F A	809
2648	TGC GGA CGA GTT TTT GAG GAG ATC TGG GAC GAC GAG ACA GTG CTC CCC	2695
810	C G R V F E E I W D D E T V L P	825
BamHI v		
2696	ATG TAC GAA GGC AGG ATC CTG GGC AAA GTG GAG AGG ATC GACTGAGCCTT	2745
826	M Y E G R I L G K V E R I D	839
2746	GGCCTCTCGCGCTGCCAACCTGGGCAAGCACCTCGCGCTGCACCATGGAGGCCAGAG	2808
2809	ACCCCTGCTCAGGCTAACAGCCACGAAATATTCTGAAGAAAAATGAAACCAATTAAAGA	2871
2872	AGACAAAGCTAGGGAGGACTGGCGCTGGCCTTACGGAGGGGGTATGTTGATCTTC	2934
2935	AGTCTCCAGGAGCCTGGTACCGAGATGAGAAAGCCTGAACATTATTAAACATGACCACT	2997
2998	CTGGGCTATAGAAGATGCTAGTGTCTTGAGAGACTGACATACATAATAGATGACTTCCTAG	3060
3061	GGTTCTGAATTATGAAACAGTCGATTCCCCCTACCCGCCCTCCCTACCCCCCGCCCCGAGTTATGC	3123
3124	GATATTATGAAACAGTCGATTCCCCCTACCCGCCCTCCCTACCCCCCGCCCCGAGTTATGC	3186
3187	TGCTTTAAACC -3'	3197

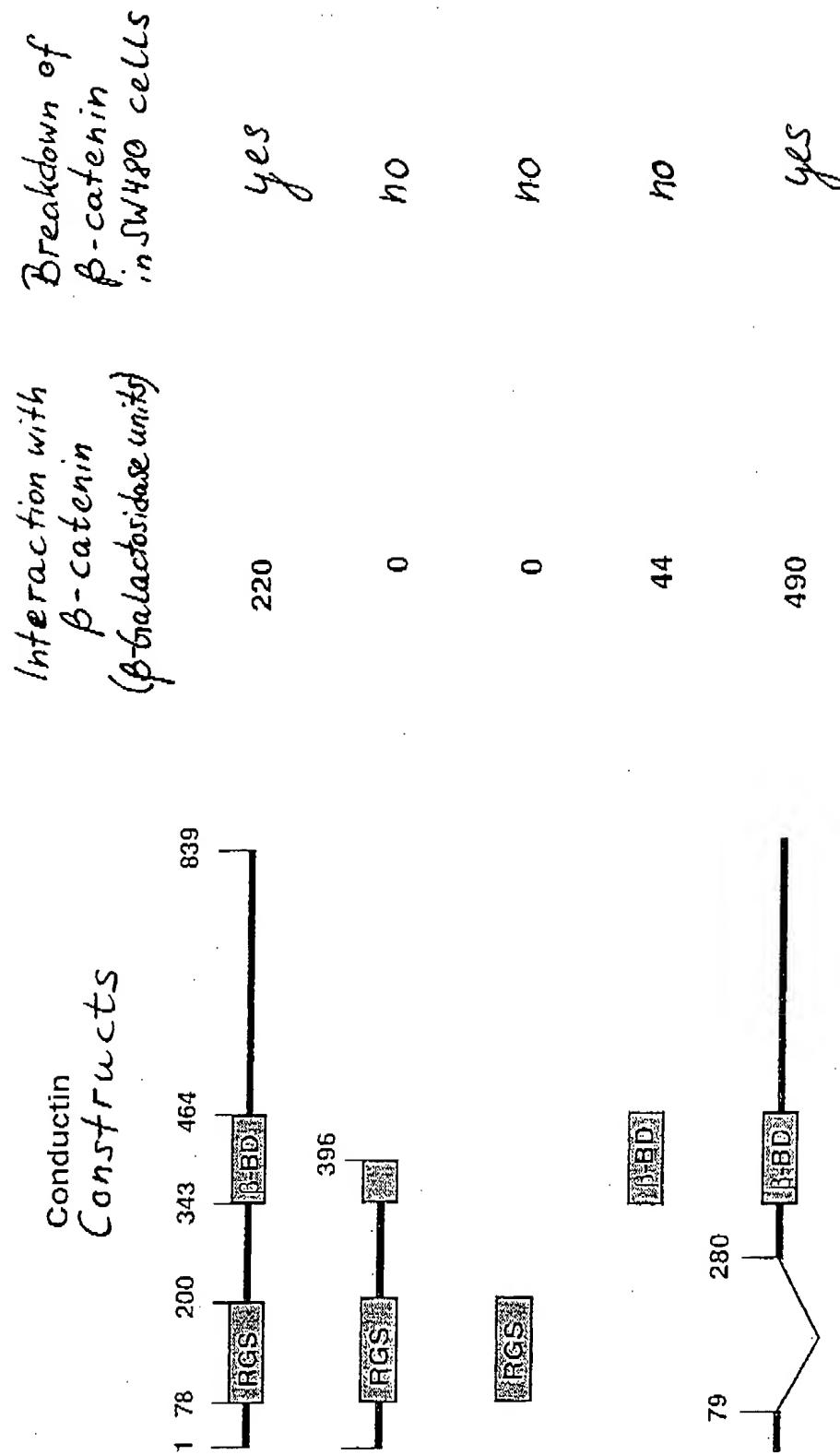
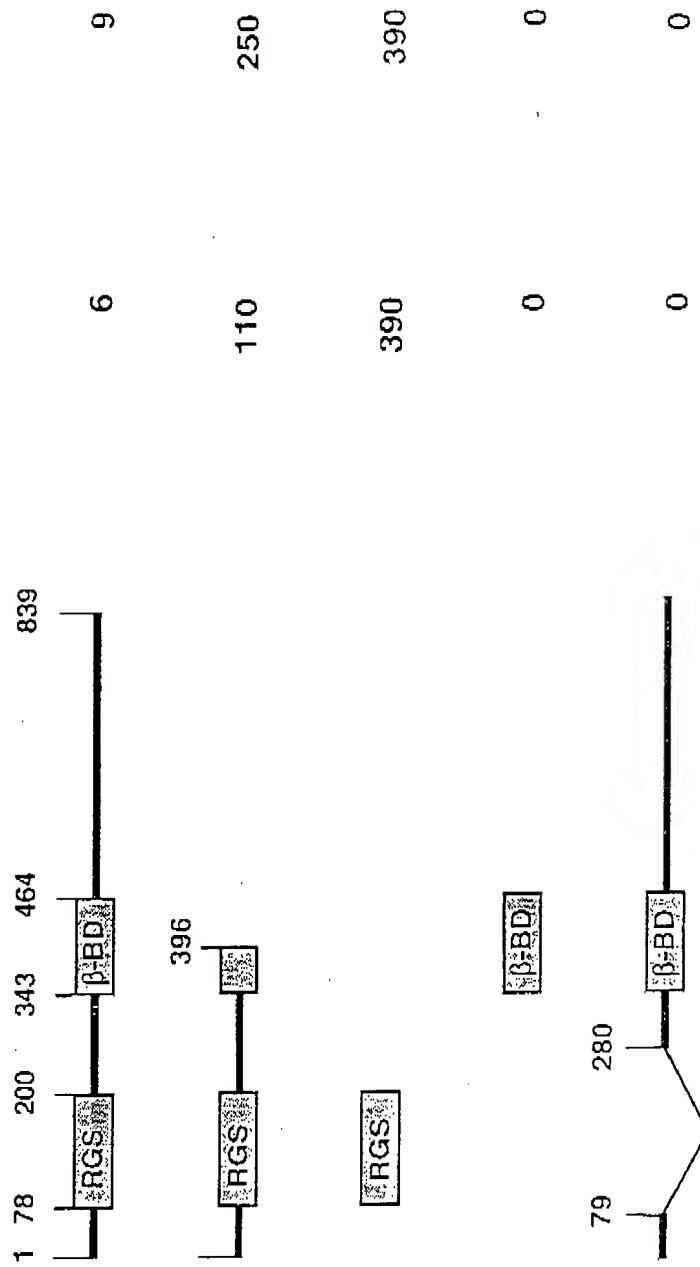


Figure 4

Conductin  
Constructs

Interaction with  
APC - fragments  
( $\beta$ -Galactosidase units)

APC Fr. 1      APC Fr. 2



17

Figure 5